Identifying and Validating New Drug Targets for Stroke and Beyond: Can Mendelian Randomization Help?

Andrew Plump, MD, PhD

George Davey Smith, MD, DSc, FRS

Until very recently, drug discovery attrition rates have been increasing.[1] This poor return on investment in biopharmaceutical R&D can be explained by scientific and extra-scientific influences. Regulatory, payer and commercial forces incentivized pharmaceutical companies to focus on common diseases and incremental improvements on existing mechanisms. The majority of resources within pharmaceutical companies have traditionally been focused on small molecule drug discovery programs, reflecting a "hammer and a nail" mindset, with small-molecule chemistry as the hammer and the nail a classically defined "druggable" target. Rather than starting with a target based on strong human biological rationale, targets have been selected based on their ability to be "drugged" by a medicinal chemist. Compounding this problem has been an inadequate understanding of disease biology and an overreliance on animal models of human disease.

Drug discovery is a long and complex process. Decisions made in discovery -- where costs are relatively modest, have major downstream effects in development, where costs increase exponentially. Today, the convergence of an innovation-friendly environment, a greater focus on translational science and the emergence of powerful new therapeutic modalities beyond synthetic small molecules has led to dramatic shifts in approach, a greater focus on transformative therapies and a gradual, but meaningful, uptick in success rates.[2] Given the long life-cycle of drug discovery, success metrics lag behind progress, however, we have begun to see a turn-around in attrition as evidenced by both apparent improvements in return on investment and United States Food and Drug Administration (FDA) approval rates.

How is innovation being driven by this intersection of extra-scientific factors, translational science and new treatment modalities? First, key stakeholders now demand innovation. Increasingly, payers will only

reimburse new medicines that are innovative and regulatory agencies have created frameworks to accelerate the development of truly innovative new medicines, such as the FDA Breakthrough Therapy, the European Medicines Agency (EMA) PRIority Medicines (PRIME) scheme, and Japan's Pharmaceuticals and Medical Devices Agency (PMDA) Sakigake designation. [3] Commercial organizations within healthcare companies fully support an innovation agenda with potential patient population size no longer solely driving commercial models. While drug pricing is a hot-button issue for society, there is general agreement that an incentive system must exist to reward innovation within the smaller rare and specialty care populations. A willingness to reimburse new therapies for smaller populations has enabled both small and large companies to focus on more homogeneous disease segments where much larger effect sizes and vastly improved therapeutic indices may be seen.

A powerful and exciting shift in the focus of drug discovery is occurring as we aspire beyond disease management to normalization and cure. This trend has been enabled by the dramatic progress in our molecular understanding of disease and the emergence of novel modalities.

Human validation of a therapeutic target, realized by the emergence of a successful first drug against such a target, is the most compelling and "safest" rationale for another drug discovery program against the same target. This explains the historical focus on next-generation -- and in many cases undifferentiated or "me-too" therapies, as we have seen repeatedly for cardiovascular diseases with classes such as the statins, angiotensin pathway modifiers and calcium channel blockers.

Human pharmacological validation for a target has limitations, whereas human genetics offers broad value and has emerged as the most compelling tool for new target identification, opening truly novel therapeutic approaches for diseases with large unmet medical need, often absent any existing treatments. A human genetic association demonstrating causality for a disease greatly increases the likelihood of success for a drug engaging the target encoded by the cognate gene, given a more direct link to causality than many other modes of investigation. Retrospective studies have proven that this is indeed the case.[4,

[5] The majority of existing drugs act on targets that have human genetic associations to Mendelian and common genetic variations, which are often identical or similar to a disease of interest. Discovery organizations in large and small biopharmaceutical companies use human genetics as a primary motivator for research investments. This is especially true for classical Mendelian diseases, which are genetically transmitted and for which we now have more than 60 FDA-approved therapies that directly target the underlying genetic anomaly.[6] This is also true for more common diseases where both Mendelian and population genetics have substantially improved our probability of success.

With the growing availability of large-scale genomic and phenotypic data over the past several years, more subtle genetic instruments have been developed. Mendelian randomization, in particular, has emerged as a powerful tool for drug discovery. Mendelian randomization provides an opportunity to interrogate nature's version of a randomized controlled clinical trial,[7-9] by using genetic variation to estimate the causal effect of an intermediate marker on clinical disease. The power of this methodology to drug discovery is best illustrated by two Mendelian randomization stories: PCSK9 and C reactive protein (CRP). Both loss- and gain-of-function genetic variation has been identified in the PCSK9 gene. These genetic variants associate strongly with circulating PCSK9 and LDL-cholesterol (LDL-C) levels, which associate strongly with coronary artery disease (CAD), consistent with a large body of epidemiological evidence. The genetics of PCSK9 was the driver of multiple drug discovery efforts aimed at identifying PCSK9 inhibitors, of which two therapies are now approved with clinically verified cardiovascular protection, exactly as predicted by the PCSK9 Mendelian randomization experiment.[10] Likewise, genetic variation in the CRP gene associates with circulating CRP levels; however, these genetic variants and their associated differences in CRP level do not relate to CAD risk. Despite a strong epidemiological association between circulating CRP levels and CAD, we conclude from the Mendelian randomization studies that CRP is not causal and thus, while a predictive biomarker of CAD, not a therapeutic drug target.[11] Today CRP is

recognized clearly as an inflammatory biomarker, prior to the Mendelian randomization studies excluding it as a therapeutic target, there were multiple industry drug discovery efforts underway.

In this edition of *Circulation*, Chong et al. publish a Mendelian randomization study designed to identify new stroke targets.[12] The authors employ a hypothesis-generating screen leveraging genome-wide association studies (GWAS) and large-scale plasma protein analysis to search for genes and putative causal biomarkers for ischemic stroke. Unlike the targeted approaches used to validate PCSK9 and invalidate CRP as therapeutic targets, this study takes a completely untargeted approach examining the genome and 653 circulating plasma proteins as possible causal factors for large artery atherosclerotic, cardioembolic and small artery occlusive stroke. The study then examines the association of "hits" for ischemic stroke subtypes with intracranial bleeding and subarachnoid hemorrhage. Hits were next evaluated in a Mendelian randomization phenome-wide association study (MR-PheWAS) across 679 diseases to understand the pleiotropic effects of these targets and to serve as proxy measures of potential safety issues for drugs modulating these targets.

For ischemic stroke, Chong et al. identified seven potentially causal biomarkers with evidence above a multiple-testing corrected threshold: five known and two novel. Among the known biomarkers were the coagulation factor, Factor XI (F11), and the apolipoprotein, lipoprotein(a) (LPA), both of which were associated with large artery atherosclerotic disease, an association that would be expected based on known biology and previous human genetic associations with other atherosclerotic diseases. While the associations with known targets do not offer new insights into drug discovery – in fact, there are ongoing drug development programs for both targets – they validate the author's approach.

The two novel observations were with respect to cardioembolic stroke (CES). The stronger of the two associations is for TWEAK (TNFS12), a circulating tumor necrosis factor (TNF)-like cytokine with a range of biological activities. Unfortunately, TWEAK had a directionally opposite association with subarachnoid

hemorrhage and multiple deleterious associations identified in the MR-PheWAS, therefore safety concerns make this a less than desirable therapeutic target.

The second novel biomarker identification was with SCARA5, a gene that encodes a scavenger receptor for Ferritin that mediates non-transferrin-dependent delivery of iron.[12] SCARA5 had a suggested protective effect for both CES and subarachnoid hemorrhage in this study. In addition, there was a trend for protection against intraventricular hemorrhage and no clear deleterious associations in the MR-PheWAS.

Is SCARA5 a compelling therapeutic target for stroke? Not now. Replication is critical for human genetic studies. Genetic variation in SCARA5 accounts for very little of the circulating SCARA5 levels and the association between SCARA5 and stroke had the lowest level of statistical evidence of the seven biomarkers. Additionally, methodological issues may affect the robustness of the conclusions. First, the authors pruned SNPs at a linkage disequilibrium r^2 of <0.10, which by no means guarantees generation of a set of "independent" variants. This means that double counting of the same signal is possible, and importantly that the different variants will have the same biases, rendering ineffective the multiple variant approaches to sensitivity analyses for horizonal pleiotropy.[13] In this situation, for example, the InSIDE assumption of MR-Egger would be violated. Second, the MR-RAPS method is not robust to unbalanced horizonal pleiotropy. In general, the statements made several times in the paper that "there was no evidence of horizontal pleiotropy" cannot be accepted on the basis of the data provided. Simply put, based on the methodology used by Chong et al., we cannot conclude that the association of the SCARA5 gene to CES is due to SCARA5 variation or that of some other factor linked to these genetic variants.

If we suspend judgment on the statistical methodology and accept the association as replicable and true, is SCARA5 then a viable drug discovery target? Still, not at this time. Human genetics offers insights into causality but ultimately the prosecution of a drug discovery program requires a deep understanding of function. What protein activity is driving disease? In what cellular compartment is this activity most

relevant? The rapidity in moving from PCSK9 genetic association to a mechanistic understanding of how PCSK9 ultimately affects LDL-C was atypical, but built from a deep understanding of cholesterol biosynthesis and LDL-receptor biology and regulation. We are not in the same place with SCARA5 biology. SCARA5 is a type II single transmembrane crossing protein receptor. Is circulating SCARA5 involved in CES causality or is it a proxy for cell surface activity? There are Mendelian randomization precedents for associations of cell surface receptors with human disease, most notably the IL6-receptor. Variation in the IL6 receptor gene influences circulating levels of IL6 and receptor-mediated effects in such a way that provides a coherent picture as to how biological activity of the pathway influences cardiovascular disease risk,[14, 15] which received subsequent support from RCT evidence.[16] A substantial body of mechanistic work is needed to better understand the function of SCARA5 and how one might modulate this target. Additionally, the broad expression of SCARA5 in the epithelium of multiple tissues suggests the potential for untoward safety issues. The MR-PheWAS suggests that low-level modulation of SCARA5 might not associate with other disease states, but what safety issues might emerge following the more potent knockdown that is likely necessary to elicit the therapeutic benefits of a drug? The identification of humans with SCARA5 deficiency could provide additional confidence in safety (and efficacy) of targeting SCARA5. We must also ask what disease we are targeting. CES is a heterogeneous disease often caused by atrial-fibrillation, so what is the subpopulation that we would target by modulating SCARA5? An understanding of disease causality is the obvious start of any drug discovery program, but an equal consideration is the targeted population and the endpoints that will be used to establish dose, proof of principle, regulatory approval and payer reimbursement.

The future of target identification clearly lies in large-scale, information-rich human genetic studies. With more sophisticated tools such as Mendelian randomization and the next-generation methodologies that build on these principles, we are enhancing our ability to identify reliable drug targets, biomarkers that can guide decision making in early drug development and patient segments most amenable to therapeutic

intervention. Translating novel findings in Mendelian randomization studies into next generation therapies will require substantial evolution in our ability to explore these newly identified targets with tools that transform them into plausible experimental agents worthy of substantial investment.

Figure 1. Human Genetics can be a powerful source for drug target identification. Mendelian randomization can provide robust instruments for new target identification and biomarker discovery. When used with a vast array of phenotypic data, Mendelian randomization (termed Mendelian randomization phenome-wide association studies or MR-PheWAS) can offer clues to the potential adverse effects of a drug against a genetically identified target. Once there is high confidence that modulation of a genetic target could benefit patients with (or at risk of) a given disease, there is still work to be done. A mechanistic understanding of disease association and a deep understanding of protein function are necessary for any drug discovery program.

ARTICLE INFORMATION

Correspondence

Andrew Plump, MD, PhD, Takeda Pharmaceutical Company Limited, 300 Massachusetts Avenue, Cambridge, MA, United States. Email andrew.plump@takeda.com

Affiliation

Andrew Plump, MD, PhD, Takeda Pharmaceutical Company Limited

George Davey Smith, MD, DSc, FRS, Medical Research Council Integrative Epidemiology Unit, University of Bristol, UK

Disclosures

George Davey Smith works in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC UU 00011/1).

REFERENCES

1. Solutions DCfH. Unlocking R&D Productivity: Measuring The Return from Pharmaceutical Inovation 2018. 2018.

2. Dowden H, Munro, J. Trends in Clinical Success Rates and Therapeutic Focus. *Nature Reviews Drug Discovery*. 2019;18.

3. Kondo H, Hata T, Ito K, Koike H and Kono N. The Current Status of Sakigake Designation in Japan, PRIME in the European Union, and Breakthrough Therapy Designation in the United States. *Ther Innov Regul Sci*. 2017;51:51-54.

4. Wang ZY and Zhang HY. Rational drug repositioning by medical genetics. *Nat Biotechnol*. 2013;31:1080-2.

5. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC and Sanseau P. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47:856-60.

6. CenterWatch. FDA Approved Drugs for Genetic Disease. 2019;2019.

7. Fisher RA. Statistical methods in genetics. *Heredity*. 1952;6:1-12.

8. Davey Smith G. Random allocation in observational data: how small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology*. 2011;22:460-3; discussion 467-8.

9. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23:R89-98.

10. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, Voros S, Giugliano RP, Davey Smith G, Fazio S and Sabatine MS. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *N Engl J Med*. 2016;375:2144-2153.

11. Collaboration CRPCHDG, Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, Shah T, Engert JC, Clarke R, Davey-Smith G, Nordestgaard BG, Saleheen D, Samani NJ, Sandhu M, Anand S, Pepys MB, Smeeth L, Whittaker J, Casas JP, Thompson SG, Hingorani AD and Danesh J. Association between C reactive

protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ*. 2011;342:d548.

12. Chong M, Sjaarda J, Pigeyre M, Mohammadi-Shemirani P, Lali R, Shoamanesh A, Gerstein HC and Pare G. Novel Drug Targets for Ischemic Stroke Identified Through Mendelian Randomization Analysis of the Blood Proteome. *Circulation*. 2019.

13. Hemani G, Bowden J and Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27:R195-R208.

14. Collaboration IRGCERF, Sarwar N, Butterworth AS, Freitag DF, Gregson J, Willeit P, Gorman DN, Gao P, Saleheen D, Rendon A, Nelson CP, Braund PS, Hall AS, Chasman DI, Tybjaerg-Hansen A, Chambers JC, Benjamin EJ, Franks PW, Clarke R, Wilde AA, Trip MD, Steri M, Witteman JC, Qi L, van der Schoot CE, de Faire U, Erdmann J, Stringham HM, Koenig W, Rader DJ, Melzer D, Reich D, Psaty BM, Kleber ME, Panagiotakos DB, Willeit J, Wennberg P, Woodward M, Adamovic S, Rimm EB, Meade TW, Gillum RF, Shaffer JA, Hofman A, Onat A, Sundstrom J, Wassertheil-Smoller S, Mellstrom D, Gallacher J, Cushman M, Tracy RP, Kauhanen J, Karlsson M, Salonen JT, Wilhelmsen L, Amouyel P, Cantin B, Best LG, Ben-Shlomo Y, Manson JE, Davey-Smith G, de Bakker PI, O'Donnell CJ, Wilson JF, Wilson AG, Assimes TL, Jansson JO, Ohlsson C, Tivesten A, Ljunggren O, Reilly MP, Hamsten A, Ingelsson E, Cambien F, Hung J, Thomas GN, Boehnke M, Schunkert H, Asselbergs FW, Kastelein JJ, Gudnason V, Salomaa V, Harris TB, Kooner JS, Allin KH, Nordestgaard BG, Hopewell JC, Goodall AH, Ridker PM, Holm H, Watkins H, Ouwehand WH, Samani NJ, Kaptoge S, Di Angelantonio E, Harari O and Danesh J. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet*. 2012;379:1205-13.

15. Interleukin-6 Receptor Mendelian Randomisation Analysis C, Swerdlow DI, Holmes MV, Kuchenbaecker KB, Engmann JE, Shah T, Sofat R, Guo Y, Chung C, Peasey A, Pfister R, Mooijaart SP, Ireland HA, Leusink M, Langenberg C, Li KW, Palmen J, Howard P, Cooper JA, Drenos F, Hardy J, Nalls MA, Li YR, Lowe G, Stewart M, Bielinski SJ, Peto J, Timpson NJ, Gallacher J, Dunlop M, Houlston R, Tomlinson I,

Tzoulaki I, Luan J, Boer JM, Forouhi NG, Onland-Moret NC, van der Schouw YT, Schnabel RB, Hubacek JA, Kubinova R, Baceviciene M, Tamosiunas A, Pajak A, Topor-Madry R, Malyutina S, Baldassarre D, Sennblad B, Tremoli E, de Faire U, Ferrucci L, Bandenelli S, Tanaka T, Meschia JF, Singleton A, Navis G, Mateo Leach I, Bakker SJ, Gansevoort RT, Ford I, Epstein SE, Burnett MS, Devaney JM, Jukema JW, Westendorp RG, Jan de Borst G, van der Graaf Y, de Jong PA, Mailand-van der Zee AH, Klungel OH, de Boer A, Doevendans PA, Stephens JW, Eaton CB, Robinson JG, Manson JE, Fowkes FG, Frayling TM, Price JF, Whincup PH, Morris RW, Lawlor DA, Smith GD, Ben-Shlomo Y, Redline S, Lange LA, Kumari M, Wareham NJ, Verschuren WM, Benjamin EJ, Whittaker JC, Hamsten A, Dudbridge F, Delaney JA, Wong A, Kuh D, Hardy R, Castillo BA, Connolly JJ, van der Harst P, Brunner EJ, Marmot MG, Wassel CL, Humphries SE, Talmud PJ, Kivimaki M, Asselbergs FW, Voevoda M, Bobak M, Pikhart H, Wilson JG, Hakonarson H, Reiner AP, Keating BJ, Sattar N, Hingorani AD and Casas JP. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet*. 2012;379:1214-24.

16. Ridker PM. Anticytokine Agents: Targeting Interleukin Signaling Pathways for the Treatment of Atherothrombosis. *Circ Res*. 2019;124:437-450.